Cloud-Recording Based Intelligent Feedback System of Voice Information

Mengwei Si, Wenwen Du, Yibo Wu, and Jiawei Chen

Shanghai Jiao Tong University, Shanghai, China
{wssg,chenjiawei}@sjtu.edu.cn, {dewwen,iceworld0324}@126.com

Abstract. In recent years, opinion mining becomes a novel and multiuse research topic. This technique has wide and many real-world applications, such as e-commerce, business-intelligence, information monitoring, public-opinion poll, e-learning, newspaper and publication compilation, business management, etc.[1] The goal we develop the Cloud-recording system is to extract information from people's communication through telephones which may be forgotten by accident. Recently, the system can be applied to automate the telephone survey while telephone surveys now are mainly done by humans. With the Cloud-recording system, companies or governments and do the surveys automatically which means telephone surveys can be at a greater amount and a lower cost. In the system, DSR is used to extract features of phones and we translate the phones to text with the help of SPHINX.

Keywords: Opinion mining, Opinion extraction, Sentiment analysis, DSR, SPHINX.

1 Introduction

Traditionally, there is no feedback from communication base station to users. People are supposed to take notes for the important information in their phone conversation. Taking notes is not convenient for two reasons. For one thing, one has to take a note book everywhere which is pretty uncomfortable. For another, there are cases in which people cannot take notes such as one is driving a car or swimming in a swimming pool. As a result, a system which can record the telephone call are very important. What's more, people will be more appreciated if the system can extract valuable information automatically. That's why the cloud-recording system is of worth.

The cloud-recording system we have developed enables the analysis of telephone survey to be automatic. Instead of hiring workers, The Company can use a computer to extract the topic and its judgment comments, evaluate whether they are positive or negative and the degree of the comments. All these analyzing process is automatically done. For example, if the respondent is asked that "What do you thing of the computer?" and the respondent answers "the computer is very fast", the cloud-recording system will detect that a respondent have assess the theme "computer" and the evaluation is "fast". At the same time, the cloud recording system will find that a respondent have make a positive evaluation become it has made a affirmative assess. Also, strength of the evaluation has been taken into account which indicates that "very fast" is a better evaluation than "fast".

Y. Wang and X. Zhang (Eds.): IOT Workshop 2012, CCIS 312, pp. 439–444, 2012.

In the system, DSR is used for extract features of phones [2]. There are two merits while using DSR. First, less data are transmitted so the costs for communication are lessened. Secondly, features can be reconstructed to wav form so that we can use SPHINX to translate the phone information into text. The SPHINX is an open source project and with which phones can be translated into text in a relatively high accurate rate [3].

After the phone information has been translated into text, text processing procedure is done. Dependency parsing is accomplished by Stanford Parser. We have design an algorithm to extract theme information and the evaluation of the theme. The system can be expressed in figure 1.

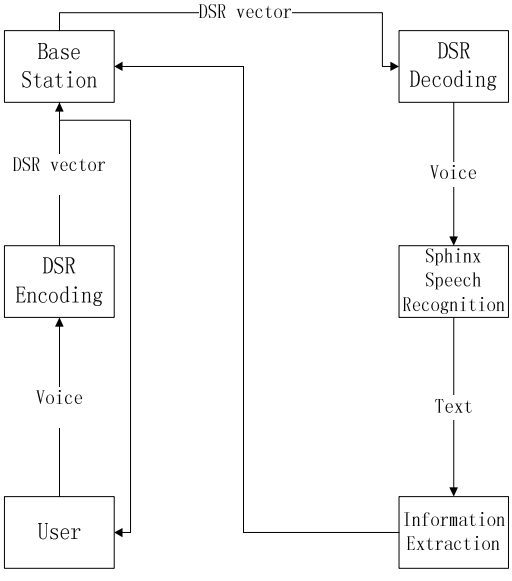

Fig. 1. System architecture

2 Related Work

2.1 Distributed Speech Recognition

A Distributed Speech Recognition (DSR) system overcomes these problems by eliminating the speech channel and instead using an error protected data channel to send a parameterized representation of the speech, which is suitable for recognition.

The value of DSR is that it provides substantial recognition performance advantages compared to a conventional mobile voice channel where both the codec compression and channel errors degrade performance. It also enables new mobile multimodal interfaces by allowing the features to be sent simultaneously to other information on a single mobile data channel such as GPRS [2].

Sphinx is an open-source speech recognition project developed by Carnegie-Mellon University since 1990 [2]. It is also the foundation of current Microsoft speech recognition technique. We use Sphinx to convert voice of wav or raw format into word document, precisely, txt format.

2.2 Sphinx

Sphinx has several different versions, among which we choose the one called Pocketsphinx.

An investigation in Sphinx reveals that it demands a user dictionary added by individual user. When analyzing the speech, Sphinx selects the most suitable words that match the voice from the dictionary, and then form a result consists of many separate words with no punctuation. In this term, it is necessary to set a specific application background,

So we can add words which are possible to appear in that condition to the dictionary. After the dictionary is founded and parameters are corrected, Sphinx can run properly under Linux and provide a reliable output.

Besides this, Sphinx also has a training section to improve its performance. It will be studied later carefully since the precision of speech recognition casts a great influence on the subsequent information extraction [3].

2.3 Stanford Parser

A natural language parser is a program that works out the grammatical structure of sentences, for instance, which groups of words go together (as "phrases") and which words are the subject or object of a verb. Probabilistic parsers use knowledge of language gained from hand-parsed sentences to try to produce the most likely analysis of new sentences. These statistical parsers still make some mistakes, but commonly work rather well [4].

3 Structure of the Automatic Telephone Survey System

We have design the application scenario of the telephone survey system as the computer ask the questions automatically and the computer wait for the respondents to answer the questions. The questions are designed by the person who wants to do a survey. And the records of the answers are transmitted to the server part to do phonetic recognition and information extraction. Then the information will be processed by the cloud recording system, the evaluation from the informant would be extracted. Then the person who wants to do the survey will get the result of the survey automatically while in the past the process has to be done by human.

4 Algorithm of Evaluation of Theme Extraction

It is designed to recognize the topic and emotional trend of the speech. The algorithm in this part includes the construction of several dictionaries (e.g. dictionary of emotional

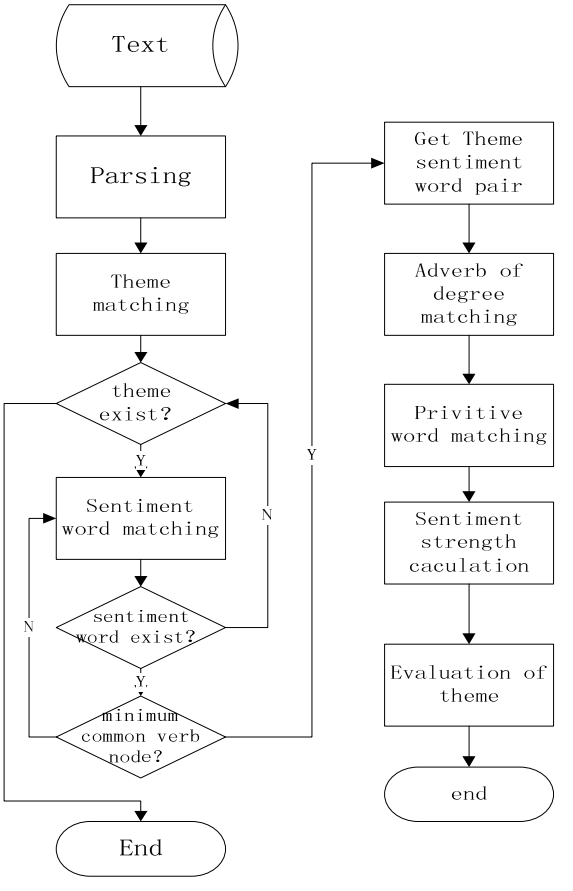

Fig. 2. Flow chart of algorithm of evaluation of sentimental strength

words, degree adverb words and topic words), matching of thematic and emotional words, dependency parsing and evaluation of the judgment degree.

In the algorithm of evaluation of theme extraction the flaw chart of which is as figure 2, when a text input the system, Stanford parser is applied to parse the text [4]. Then themes and sentimental words will be matched from the dictionary of them. Sentimental words and themes which shared the same minimum common verb node will a theme-sentiment pair. Then the sentimental strength of the sentiment word will be calculated. In order to calculate the strength, adverb of degree and privatives should be taken into account. Then, after considering all the conditions, the algorithm will give the sentimental strength of a theme. Then, sentimental strength of other themes will be calculated in the same way. Hence, we get the evaluation of all the themes in the text.

The following is some explanation of some of the key points.

First, we need to construct them dictionary, sentimental word dictionary, privative dictionary and adverb of degree dictionary. As we have constrained in the area of telephone survey of computer products. The theme dictionary is constructed in the area of computer. For example, computer, keyboard, performance are all in the theme

dictionary. Sentimental word dictionary and adverb of degree dictionary are from Hownet [5]. Some post processes have been done to make the dictionaries suitable to our application. For instance, some of the sentimental words are positive in some areas while negative in other areas. The word "fast" is positive in the sentence "The computer is fast" while "The driver drives too fast" conveying a negative opinion. Thus, we have to redesign the dictionary manually. This process has been done partly. The reasons are twofold. For one thing, the amount of work is too large. For another, it is not necessary to do the work for the influence of small part of sentimental word may not have too much damage to the performance of the system. Privative words dictionary are constructed manually for the reason that the number of privatives is relatively small. As to adverb of degree dictionary, we have to reclassify the words according to our algorithm.

Secondly, the Stanford parser serves as a tool to parse the sentences. For the output of the parser, we can get the pos tagging and parsing of a sentence and the parser provide us a tree of parsing result. We have found that a sentiment word tend to qualify the theme which share the same minimum parent verb node. In the algorithm, we design a method to find the minimum parent verb node. As a result, the algorithm detects the verb and determines the theme-sentiment pair. In the theme-sentiment pair, the sentiment word qualifies the theme. It should be noted that a theme can be qualified by numbers of sentiment words but a sentiment word can only qualify a single theme.

Thirdly, sentiment strength of the evaluation needs to be defined before we determine the sentiment strength of the evaluation of a certain topic [5]. Real number is used to determine the sentiment strength of an evaluation. For example, the sentence "The computer is fast", sentiment strength of the evaluation is 1 while the sentence "The computer is very fast", sentiment strength of the evaluation is 2.

Fourthly, adverb of degree enhance the sentimental strength while privatives adverse the sentimental strength. The combine of the two kinds of word sometimes entrance while sometimes weaken the sentimental strength. So we have to classify the adverb of degree carefully to make it distinct. For example, in the sentence "The computer is very fast" the sentiment strength of the evaluation is 2. However, in the sentence "The computer is not very fast" the sentiment strength of the evaluation will never be -2 because the privative waken the sentimental strength [7]. But in fact sometimes, the privative can enhance the sentimental strength such as the sentence "The computer is not a bit fast" [8].

The performance of the algorithm when test in pure text as show in table 1. From the table we can see if the input is pure text the performance of the system is acceptable. However, it is one of the defects of the system that when real phone calls are put into the system, the error rate will increase markedly and we are trying hard to reduce the differences.

Table 1. Performance of Algorithm of evaluation of theme extraction

Theme recall rate	Accurate rate of evaluation polarity	F-measure
89.8%	62.9%	76.3%

Acknowledgments. This work gets a lot of support from the MediaSoc Lab of Shanghai Jiao Tong University. Here we would like to express our heartfelt thanks to MediaSoc for all its help throughout the work.

References

1. Yao, T.F., Nie, Q.Y., Li, J.C., et al.: An Opinion Mining System for Chinese Automobile Reviews. In: Proc. of Chinese Information Processing Advanced Progress – Chinese Information Federation 25th Anniversary Conference, pp. 260–281. Press of Tsinghua University (2006)
2. http://www.etsi.org/WebSite/Technologies/DistributedSpeechRecognition.aspx
3. http://cmusphinx.sourceforge.net/
4. http://nlp.stanford.edu/software/lex-parser.shtml
5. Dong, Z., Dong, Q.: Hownet, http://www.keenage.com
6. Kim, S.-M., Hovy, E.: Determining the Sentiment of Opinions. In: Proceedings of COLING 2004, the Conference on Computational Linguistics (COLING 2004), Geneva, Switzerland, pp. 1367–1373 (2004)
7. Yao, T., Lou, D., Fang, X.: Polarity Distinction for Chinese Sentiment Words ICCC 2007 paper (IC240)
8. Si, M.W., et al.: Advanced Materials Research 808, 217–218 (2011)